



Clustering on Customer Dataset

By Salih Eren
Yüzbaşıoğlu
AI Club AR-GE

CUSTOMER IDENTITY ANALYSIS

Customer identity analysis is a crucial aspect of modern marketing strategies that enables companies to gain valuable insights into their customer base. It involves collecting and analyzing data about customers' demographic information, purchasing behavior, preferences, and interactions with the brand. By understanding customer identity, companies can tailor their marketing efforts, enhance customer experiences, and ultimately increase profitability.

Here are some key reasons why customer identity analysis is important:

Personalized Marketing: By analyzing customer identity, companies can create personalized marketing campaigns that resonate with individual customers. This targeted approach leads to higher engagement, conversion rates, and customer loyalty. For example, Amazon utilizes customer identity analysis to provide personalized product recommendations, resulting in increased sales and customer satisfaction.

Customer Segmentation: Customer identity analysis helps in segmenting customers into distinct groups based on similarities in their characteristics and behavior. This segmentation allows companies to develop specific strategies for each segment, optimizing marketing efforts and resource allocation. Nike successfully implemented customer identity analysis to identify different customer segments, leading to the creation of specialized products and targeted marketing campaigns.

Improved Customer Experience: Understanding customer identity enables companies to deliver exceptional customer experiences. By analyzing customer preferences, companies can provide personalized services, tailored offers, and relevant content. This results in increased customer satisfaction, brand loyalty, and positive word-of-mouth. Netflix effectively utilizes customer identity analysis to recommend personalized content, leading to higher user retention and engagement.

Retention and Churn Prevention: Customer identity analysis helps in identifying potential churn risks and implementing targeted retention strategies. By monitoring customer behavior patterns and preferences, companies can proactively address customer concerns, offer incentives, and ensure long-term customer loyalty. Spotify leverages customer identity analysis to understand user preferences, curate personalized playlists, and prevent churn, resulting in increased subscription renewals.

Product Development and Innovation: Analyzing customer identity provides valuable insights into customer needs, preferences, and pain points. Companies can utilize this information to develop new products, enhance existing offerings, and stay ahead of competitors. Apple's customer identity analysis played a pivotal role in the development of the iPhone, as they identified the demand for a multifunctional mobile device that catered to various customer segments.

Companies that have effectively implemented customer identity analysis and experienced increased profits include:

Amazon: Through personalized recommendations and targeted marketing, Amazon has significantly increased its sales and customer retention.

Nike: Nike's customer identity analysis has allowed them to create specialized products and targeted marketing campaigns, resulting in increased revenue and brand loyalty.

Netflix: By leveraging customer identity analysis to recommend personalized content, Netflix has achieved higher user retention and engagement, leading to substantial profit growth.

Spotify: Spotify's analysis of customer identity has allowed them to curate personalized playlists and prevent churn, resulting in increased subscription renewals and profitability.

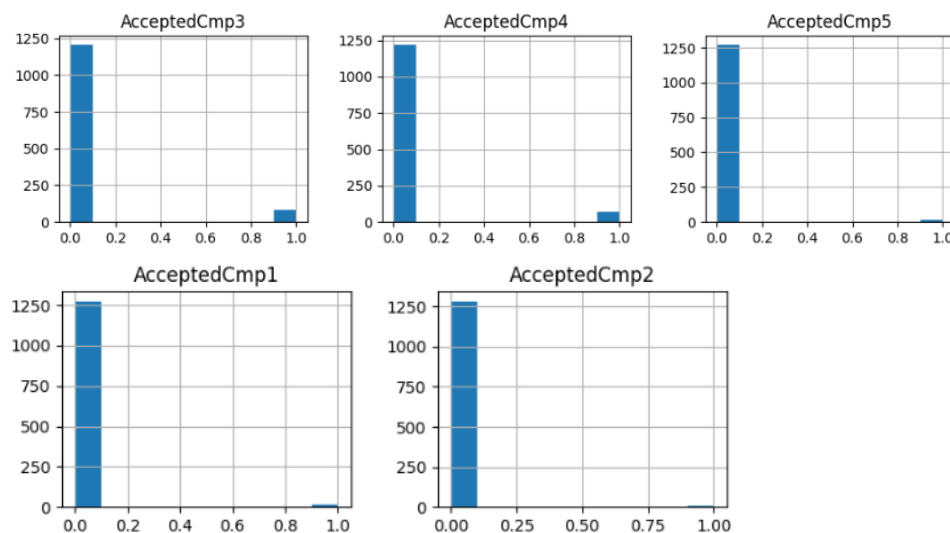
Apple: Apple's understanding of customer identity led to the development of the iPhone, which revolutionized the smartphone industry and contributed to Apple's substantial profit growth.

EXPLORING DATA

We are analyzing a dataset that includes information on over 2000 individuals. Each person's profile comprises 28 distinct attributes, such as age, income, education, and purchasing habits. This dataset provides valuable insights into individual preferences, including their buying patterns and responses to marketing campaigns. Additionally, we possess supplementary data on their store, catalog, and web-based purchases, as well as their participation in campaigns. By employing cutting-edge methodologies, our objective is to identify target demographics for campaigns and optimize advertising strategies accordingly.

Campaigns

Looking at campaigns only a small number of people use them so changing the campaigns or removing them whole would be a good idea. Below are campaign graphs.



It is obvious these campaigns don't work well.

Correlations

Correlation is a statistical concept that holds significant relevance for business professionals and decision-makers. It refers to the statistical relationship or association between two or more variables. Understanding correlation enables business individuals to comprehend how changes in one variable correspond to changes in another variable.

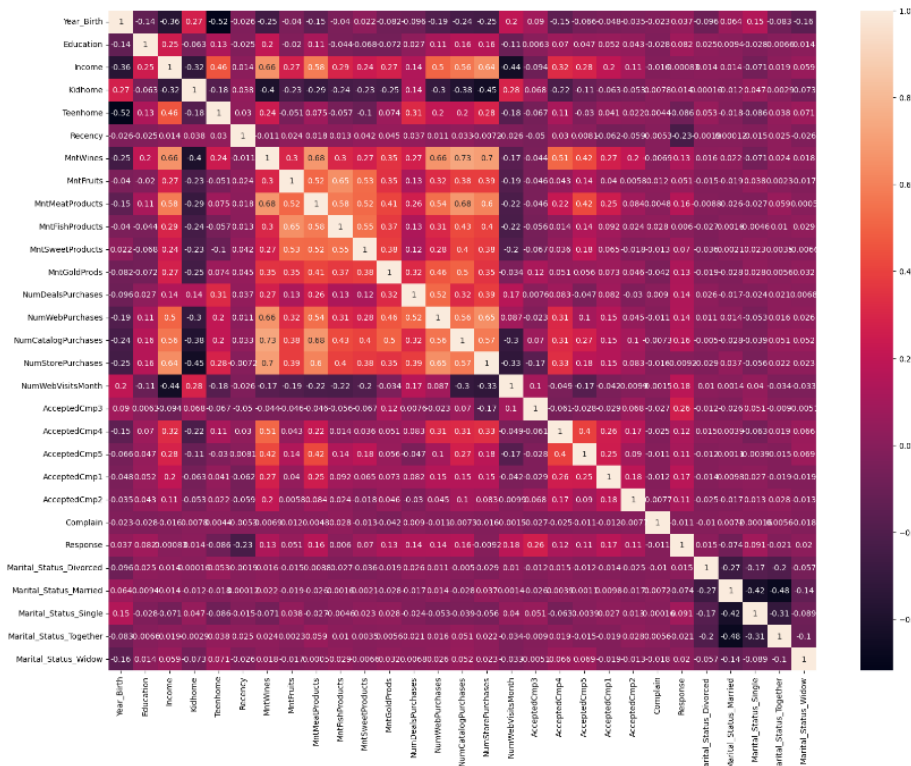
In the business world, correlations can provide valuable insights into various aspects of operations, customer behavior, market trends, and performance indicators. By examining the correlation between different variables, businesses can identify patterns, make informed predictions, and make strategic decisions.

So, let's look at some of the correlations in your data:

Upon analyzing the data, we observe strong correlations between various products in terms of consumer purchasing patterns. It appears that individuals who purchase meat also tend to buy wine, while those who acquire wine often exhibit a preference for fruit. Similarly, customers who purchase fish frequently show a propensity for sweet products. Moreover, we find a correlation between web visits and purchases made through promotional deals, whereas catalog purchases exhibit a correlation with the acquisition of gold products. These findings suggest a general association

between income, meat, wine, and overall purchasing behavior. Additionally, it is worth noting that individuals who responded positively to campaign 4 demonstrate a particular affinity for wines.

Below figure shows correlations in its pure form.



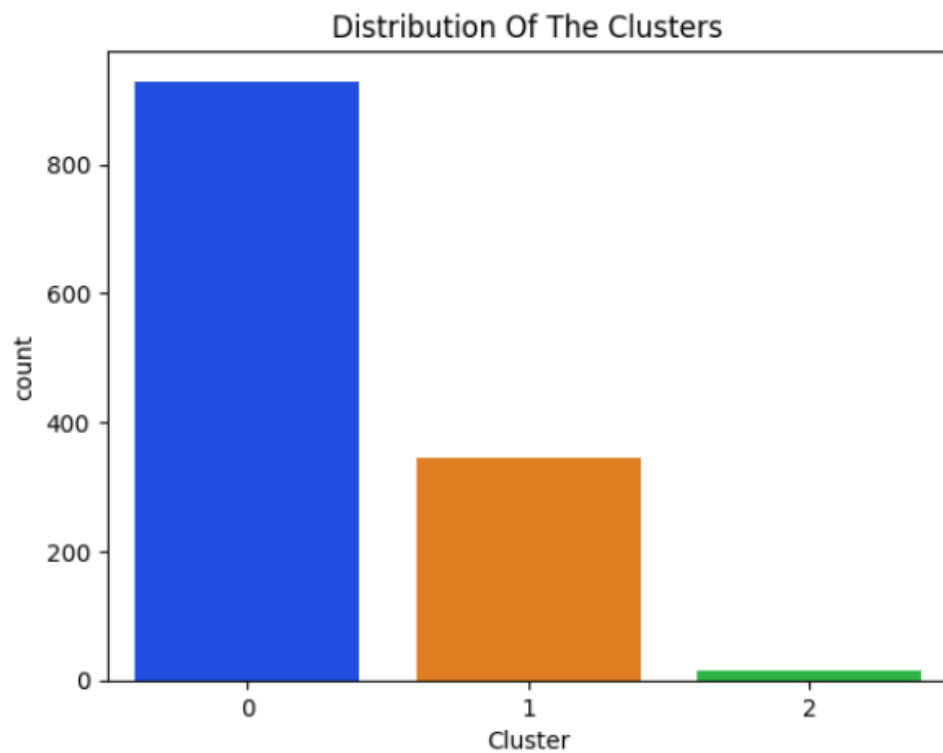
Clustering

Clustering is a valuable concept in the realm of business that can provide significant insights and benefits for decision-making. In simple terms, clustering refers to a process of grouping or categorizing similar objects or data points together based on their shared characteristics or attributes.

For a business professional, clustering can be a powerful tool for understanding customer segments, market trends, and patterns within their data. By applying clustering techniques to customer data, businesses can identify distinct groups or clusters of customers who exhibit similar behaviors, preferences, or characteristics. This process enables businesses to gain a deeper understanding of their customer base and tailor their strategies accordingly.

We have trained extensive and state-of-art machine learning models to cluster the data into 3 groups. We will call these groups blue, orange and yellow.

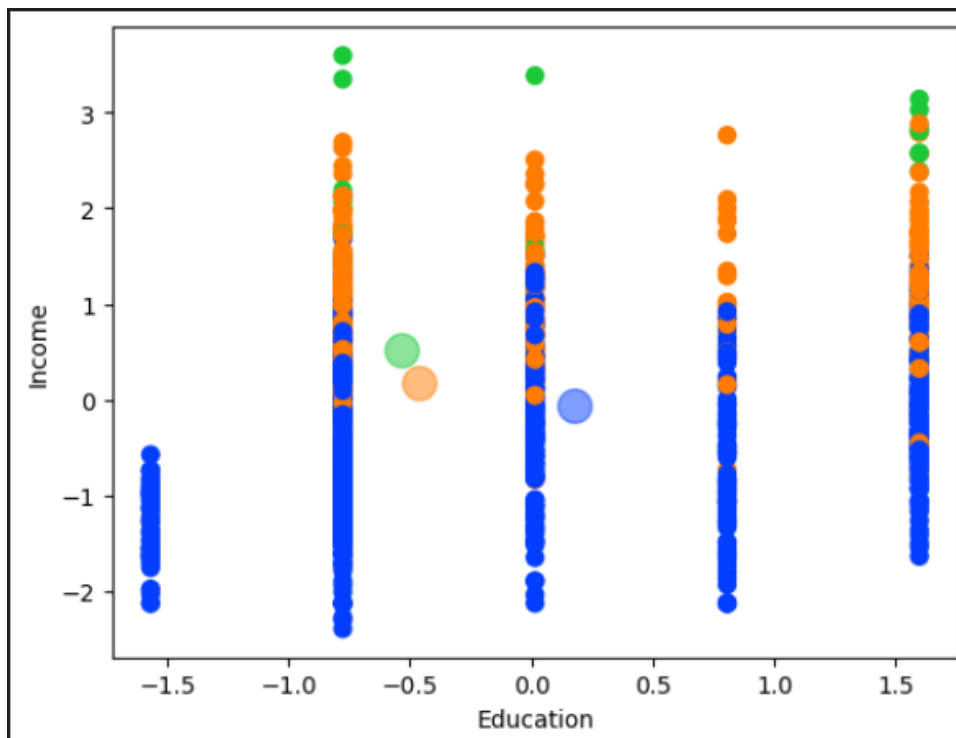
Blue has the most people while orange has less and yellow the smallest. Below figure represents their people count.



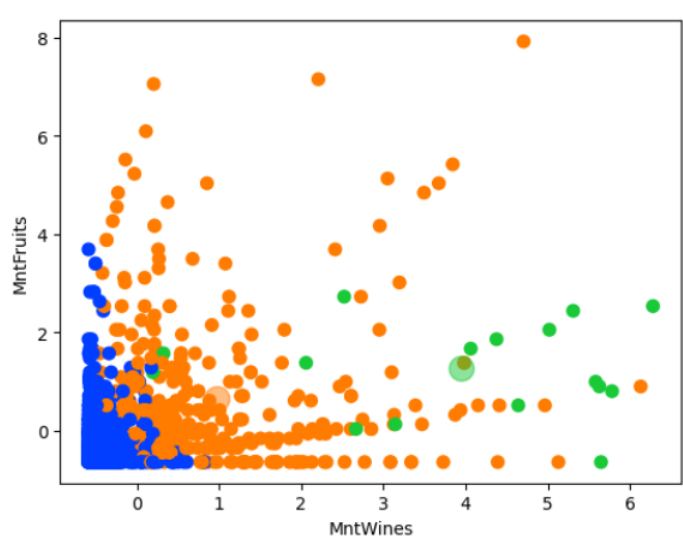
Insight to Data and Advice

Looking at cluster's Income and Education we see below graph.

Every node represents a human, and a big node represents the means of every group. Also 0 means mean in this type of figure since data is normalized. Also, every color represents a cluster.



Below figure shows fruit purchases and wine purchases colored such that every node belongs to a cluster.



Since we so much data instead of showing each one of them. We analyzed every feature in the data, and we created a table to show features of each data and combined them to make the table below.

Blue	Orange	Green
Average income and Education	High income and education	Highest income& education

Small spent on wine & fruits	More on wine and fruits	So much on wine and some in fruits
Young and lots of kids	Old, not so many kids	Old, very small amount of kids
Did not accept any campaign	Did not accept any campaign	Accepted campaigns 1,2,4,5
Small number of teens at home	High number of teens at home	Small number of teens and high response rate
Average sweet and gold purchase	Lots of gold and some sweets	Lots of sweets and some gold
Does not use web or deal purchases	Uses web and deal purchases	Some web purchases but none deal purchases
Does not use catalog or store	Some catalog and store purchases	Lots of catalog and store purchases
Normal amount of web visits	Normal amount of web visits	Does not use web often

Looking at the table we see only green people accepted campaigns and they had a high response rate. So, it would make sense to make campaigns targeting green people.

Green individuals primarily comprise individuals with high income, education, and advanced age. They exhibit a strong affinity for indulging in wine and sweets, alongside occasional purchases of fruits and gold. Their preferred mode of shopping predominantly revolves around catalog or in-store purchases, with limited utilization of online platforms.

Given their inclination towards participating in campaigns, it is advisable to design targeted initiatives tailored specifically to engage this demographic. To maximize the impact of such campaigns, it would be beneficial to align them with the interests of green individuals, particularly focusing on themes related to wine and sweets. Additionally, ensuring prominent visibility of these campaigns in catalogs and physical stores would yield optimal outcomes.

About blue people, they have an average income and education, they are young, and they have kids, and they don't have any affinity towards any product, and they compose most of the individuals.

Since their response rate and campaign acceptance is small targeting them in a specific way would not be a good idea. Instead, since there are a lot of kids, we can put products targeted for kids inside the market. Like school products or games. This group also buys some sweets and gold. So, putting sweets in a very broad place inside the market would be a good idea to make them buy sweets.

Orange individuals constitute a significant portion of the population, surpassing the green demographic but falling short of the blue cohort. They possess considerable affluence and educational attainment. While they are generally older and have fewer children, they do have teenagers within their households. Notably, orange individuals stand out as the sole group that frequently engages in online purchases. Their shopping preferences encompass substantial acquisitions of gold, alongside

moderate indulgence in sweets, wine, and fruits. Similar to the blue category, they exhibit resistance to campaign initiatives.

To effectively target orange individuals, it is advisable to promote gold, sweets, wine, and fruits through online advertising, considering their propensity for web-based transactions. Additionally, targeting their teenagers with advertisements featuring products tailored to their interests can prove fruitful.

Moreover, the incidence of customer complaints is relatively low, which could indicate a satisfactory level of customer happiness or potentially reflect the challenges associated with lodging complaints. Consequently, it would be prudent to develop enhanced mechanisms for addressing customer grievances. One viable approach could involve incorporating a dedicated complaint page on the website.